



## Intitulé de l'offre de stage

Feature engineering et Deep Learning sur images microscopiques pour la classification moléculaire de cancer du sein

<b>Stagiaire en</b>	Informatique (Deep Learning / Classification)
<b>Affectation</b>	Équipe commune CRCT/IRIT
<b>Durée</b>	5 à 6 mois
<b>Rémunération</b>	environ 1230 € net mensuel
<b>Date de la publication</b>	01/01/2022
<b>Date d'embauche prévue</b>	mars ou avril 2022 suivant disponibilités
<b>Lieu</b>	CRCT, 2 Avenue Hubert Curien – 31100 Toulouse

## Le Centre de Recherche en Cancérologie de Toulouse

Le CRCT est une unité de recherche conjointe entre l'Inserm et l'Université Toulouse III Paul Sabatier. Au cœur de l'Oncopole de Toulouse, le CRCT, avec tous ses partenaires (institutionnels, universitaires, cliniques, industriels, caritatifs...), stimule l'innovation en termes de recherche et d'enseignement dans la lutte contre le cancer. Le CRCT conduit une approche intégrée entre la recherche, les soins et l'enseignement, dans une logique transversale et multidisciplinaire.

## L'Institut de Recherche en Informatique de Toulouse

L'Institut de Recherche en Informatique de Toulouse (IRIT), une des plus imposantes Unité Mixte de Recherche au niveau national, est l'un des piliers de la recherche en Occitanie avec ses 700 membres, permanents et non-permanents. De par son caractère multi-tutelle (CNRS, Universités toulousaines), son impact scientifique et ses interactions avec les autres domaines, le laboratoire constitue une des forces structurantes du paysage de l'informatique et de ses applications dans le monde du numérique, tant au niveau régional que national.

## L'équipe

L'informatique et l'intelligence artificielle sont appelés à prendre une place croissante dans le monde de la recherche médicale, et en particulier dans le monde de la recherche contre le cancer. L'équipe commune IRIT/CRCT, co-localisée sur le site de l'Oncopole de Toulouse, a pour but de faire travailler ensemble et sur le même site chercheurs en informatique, chercheurs dans le domaine du cancer et médecins.

## Objet du stage

Le diagnostic des cancers du sein est réalisé par le pathologiste à partir de l'analyse microscopique du tissu tumoral. Cela permet de définir d'une part l'agressivité des tumeurs et d'autre part la prise en charge thérapeutique initiale des patientes. Cependant, il est fréquent que la simple analyse microscopique soit renforcée par une analyse génomique permettant de mieux évaluer le pronostic du cancer et d'en donner sa classification moléculaire. Cette dernière pourrait être approchée par une analyse extensive des caractéristiques microscopiques des noyaux des cellules tumorales sans recourir à l'analyse génomique. Dans ce cadre, l'équipe d'anatomo-pathologie développe en collaboration avec l'IRIT un projet permettant d'annoter parfaitement un dataset de lames en coloration standard avec les structures caractéristiques des tumeurs ; en effet, alors que ce travail est généralement fait manuellement par détournage par les pathologistes sur les images (avec tous les problèmes d'oublis et/ou d'erreurs afférents), nous avons développé une technique basée sur l'immunohistochimie qui nous permet une annotation automatique et parfaite du dataset. A partir de ce dataset, il sera possible de réaliser un apprentissage de type Deep Learning afin d'extraire automatiquement ces "features" des images en coloration standard, puis de réaliser une classification automatique basée sur ces features.

Dans ce cadre, les objectifs du stage sont :

- > de réaliser une étude bibliographique afin d'avoir un état de l'art en matière d'extraction de features sur les noyaux des cellules tumorales, que ce soit à partir d'approches classiques (morphologie, texture) ou d'approches Deep Learning ;
- > d'appliquer ces méthodes pour extraire des caractéristiques nucléaires à partir d'un dataset d'images microscopiques de cancers du sein ;
- > de tester des algorithmes de classification pour évaluer le type moléculaire de la tumeur à partir des features extraites sur ces images, avec une attention particulière portée à l'interprétabilité de ces méthodes (XAI, eXplainable Artificial Intelligence).

Un communiqué de presse détaillant les objectifs de ce projet porté par l'IUCT-Oncopole est disponible [ici](#).

## Formation

École d'ingénieur, de préférence avec spécialisation en informatique ou mathématiques (en année de césure ou stage long). Master 2 informatique ou mathématiques appliquées.

## Compétences attendues

- > Une bonne maîtrise du langage de programmation python ;
- > La connaissance des bases du Machine Learning et du Deep Learning (en particulier sur images) ;
- > La connaissance d'un framework Deep Learning (pytorch, tensorflow) sera appréciée ;
- > La connaissance d'algorithmes classiques de vision par ordinateur sera appréciée.

## Modalité de candidature

- > CV à envoyer à: [apriorics@stages-medecine-numerique.fr](mailto:apriorics@stages-medecine-numerique.fr)
- > Date limite de candidature : 31/01/2022
- > Encadrants: Dr Camille Franchet (IUCT-O), Robin Schwob (Thales), Sandrine Mouysset (IRIT), Daniel Ruiz (IRIT), David Simoncini (IRIT), Jean-Marc Alliot

Une première sélection sera effectuée sur la base des CVs reçus. Les candidats seront immédiatement informés du résultat, et ceux qui seront retenus à l'issue de la première sélection seront invités à un ou deux entretiens individuels en téléconférence avant sélection définitive.

L'ensemble des offres de stage est disponible sur <http://www.stages-medecine-numerique.fr>.